An improved method to search for flares from point sources of ultra-high-energy photons

Jarosław Stasielak[1], Nataliia Borodai[1], Dariusz Góra[1] and Marcus Niechciol[2]

[1]*Institute of Nuclear Physics PAN, Kraków, Poland*
[2]*Department of Physics, University of Siegen, Siegen, Germany*

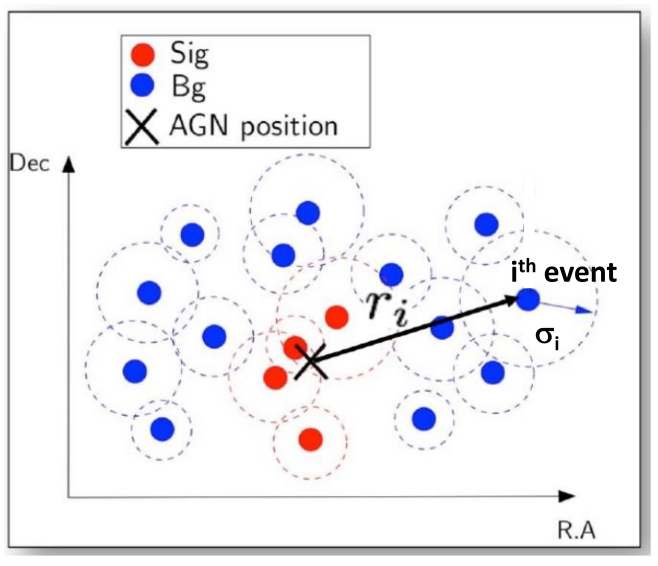
Take-home messages

- **Identification of clustering in cosmic-ray data would provide evidence for possible existence of ultra-high-energy (UHE) photons (E > 10^{17} eV) and could potentially help identify their sources.**
- **The presented stacking method, which uses a time-clustering algorithm combined with an unbinned likelihood study, is able to distinguish between events initiated by photons and those initiated by hadrons (background).**
- **The stacking method with a photon tag requires only a few events to identify a photon flare.**

Motivation

- **Astrophysical flares may be the source of some cosmic rays, which, if they are photons, should group into clusters of events correlated in space and time.**
- **Identification of such clustering in data would provide important evidence for the existence of UHE photons, i.e. with energy > 10^{17} eV.**
- **Search for space-time clustering**
  - **might be effective to search for the cosmic ray sources**
  - **can help to put stronger limits on the flux of UHE photons.**

σ_i - the angular uncertainty of event
r_i - the angular distance of event from source



- To identify flare/flares from a point source, we have to find an excess of events from a particular direction over the background.

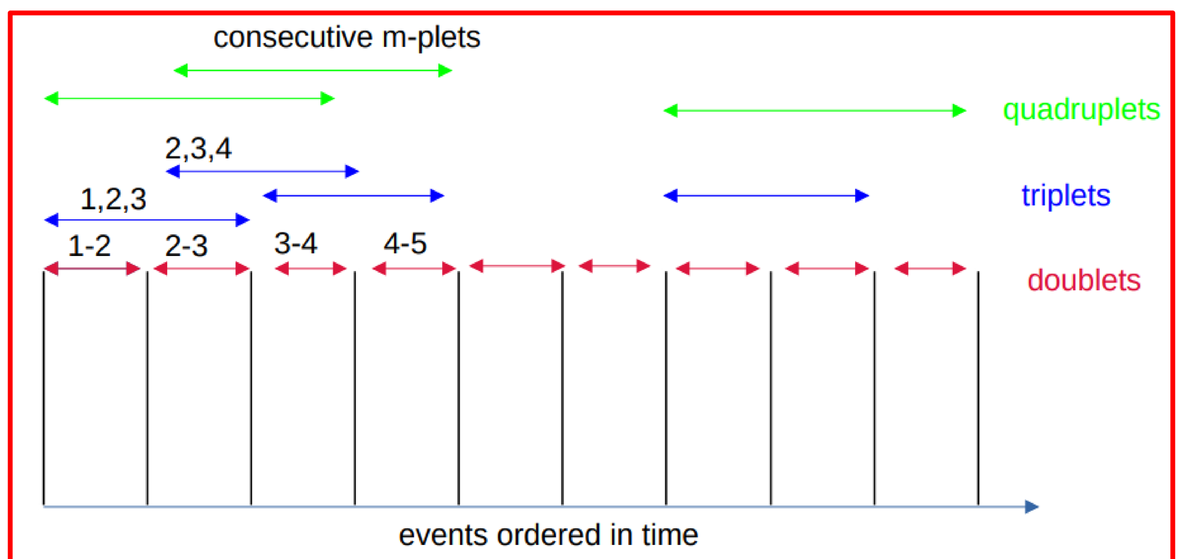
Standard space-time clustering analysis

- J. Braun et al. Astropart. Phys. 29 (2008) 299 + time search
- Maximize the likelihood of possible multiplets in a data sample (doublets, triplets, quadruplets, etc.) and calculate test statistic:
  - likelihood that flare consists of **n** signal events within a given multiplet time window ΔT_j

$$TS_j(n) = -2\log\left(\mathcal{L}(0)/\mathcal{L}(n)\right)$$

  - the test statistic

$$\mathcal{L}(n) = \prod_{i=1}^{N}\left(\frac{n}{N}s_i + (1-\frac{n}{N})b_i\right)$$



Combined signal PDF $s_i = s_i^{space} s_i^{time}$

Gaussian spatial PDF $s_i^{space} = \frac{1}{2\pi\sigma_i^2}\exp\left(-\frac{|\vec{r}_i - \vec{r}_s|^2}{2\sigma_i^2}\right)$

Heaviside temporal PDF $s_i^{time} = \frac{H(t_j^{max} - t_i)H(t_i - t_j^{min})}{\Delta t_j}$

Background PDF $b_i = b_i^{space} b_i^{time}$

$b_i^{space} = 1/\Delta\Omega$

$b_i^{time} = 1/\Delta t_{data}$

time span of a given multiplet

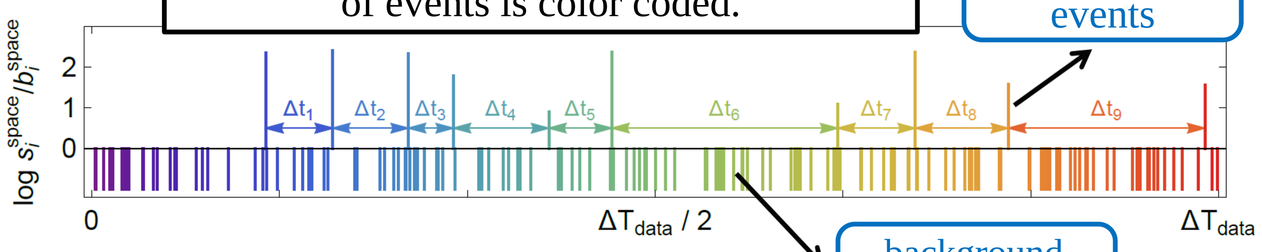
N - number of all events
σ_i - angular reconstruction uncertainty of event i
$\vec{r}_i$, $\vec{r}_s$ - direction to event i and source
$\Delta\Omega$ - solid angle

- We obtain estimates of:
  - number of signal events (n_s)
  - the flare duration (ΔT), i.e. time span ΔT_j of the most significant multiplet (multiplet with the highest TS_j)

Improved method of space-time clustering analysis

- **Stacking method** (D. Góra et al. Astropart. Phys. 35 (2011) 201) **use only doublets, consists of 3 steps.**
- Select flare candidates from the data using solely space information: identify signal-like events based on the ratio of the signal PDF to background PDF > threshold S/B, i.e. s_i^{space}/b_i^{space} > S/B **1st step**

Data collected over time period ΔT_{data}. Time of events is color coded.



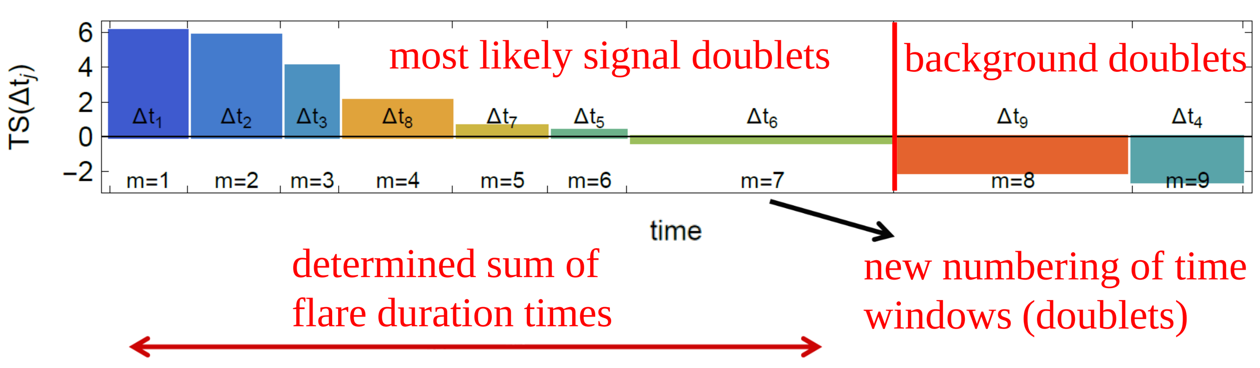
- Extract all consecutive doublets to isolate all possible time windows Δt_j that compose the flares contribution.

- For each doublet j maximize test statistic $TS_{\Delta t_j}(n)$ (calculate doublet significance) using standard method with marginalization term to provide more uniform exposure for finding doublets of different widths:

$$TS_{\Delta t_j}(n) = -2\log\left[\frac{\Delta T_{data}}{\Delta t_j}\mathcal{L}(0)/\mathcal{L}(n)\right]$$ **2nd step**

- Only events within Δt_j interval are taken into account, thus in this step both space and time information is used.
- Sort doublets according to the value of $TS_{\Delta t_j}$, i.e. to their significance, and change numbering of doublets introducing multiplicity index **m**.

How to separate signal doublets from the background?



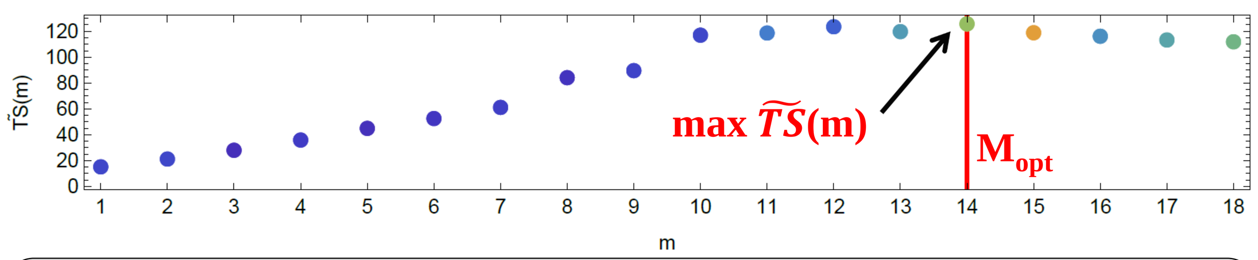
- **Stacking analysis:**
  - one-event signal PDF s_i is replaced by the weighted sum of signal sub-terms over m doublets, where weights $w_j = TS(\Delta t_j)$

$$s_i \rightarrow s_i^{tot}(m) = \Sigma_{j=1}^{m} w_j s_i^j / \Sigma_{j=1}^{m} w_j$$ **3rd step**

  - use standard likelihood and test statistic with stacking term $s_i^{tot}(m)$:

$$\mathcal{L}(n) \rightarrow \mathcal{L}(n,m)$$
$$TS \rightarrow \widetilde{TS}(m) = -2\log\left[\mathcal{L}(0)/\mathcal{L}(n,m)\right]$$

- Maximize $\widetilde{TS}(m)$ to find optimal (total) number of doublets in all flares (M_{opt})



- M_{opt} determines **total flares duration**. It is **estimated as a sum of most significant, thus not necessarily consecutive** doublets: $\Delta T = \Sigma_{m=1}^{M_{opt}} \Delta t_m$

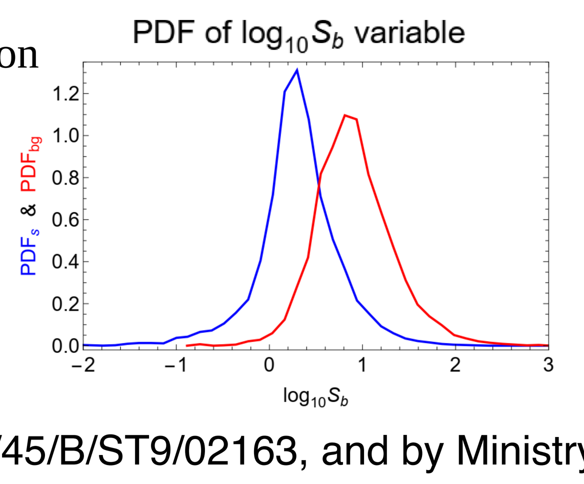
Application of the S_b photon tag to enhance sensitivity for photons search

- S_b variable is used to discriminate between photons and (background) showers

$$S_b = \Sigma_{i=1}^{n} S_i\left(\frac{R_i}{1000\ m}\right)^4$$

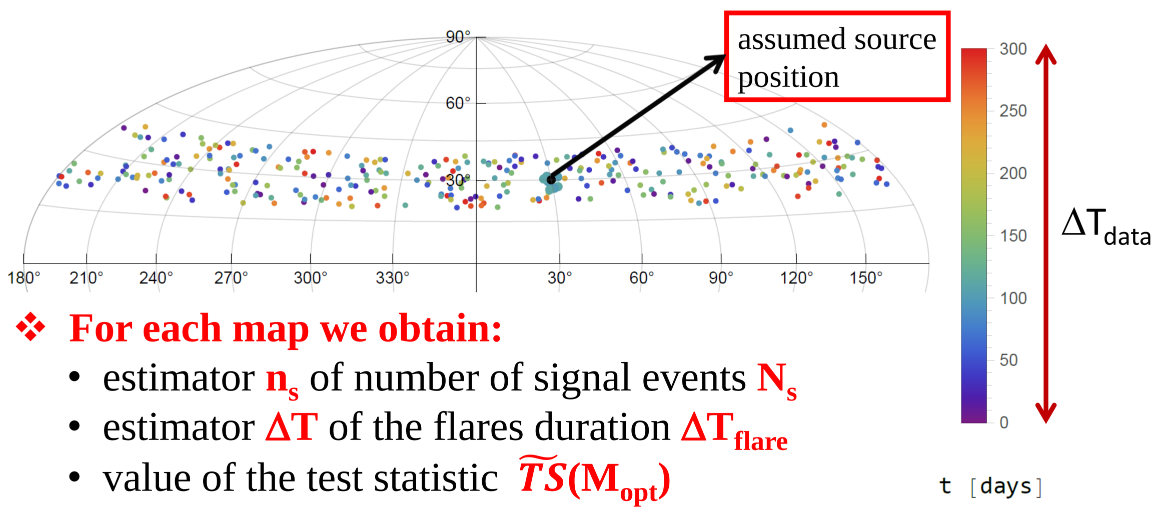
detector signal; distance from the shower axis

- **Photon tag:** Probability Distribution Functions of $\log_{10}S_b$ for **photons** (PDF_s) and **protons** (background) (PDF_{bg}), make the replacement
  - s_i^{space} => $s_i^{space} * PDF_s(S_b)$
  - b_i^{space} => $b_i^{space} * PDF_{bg}(S_b)$



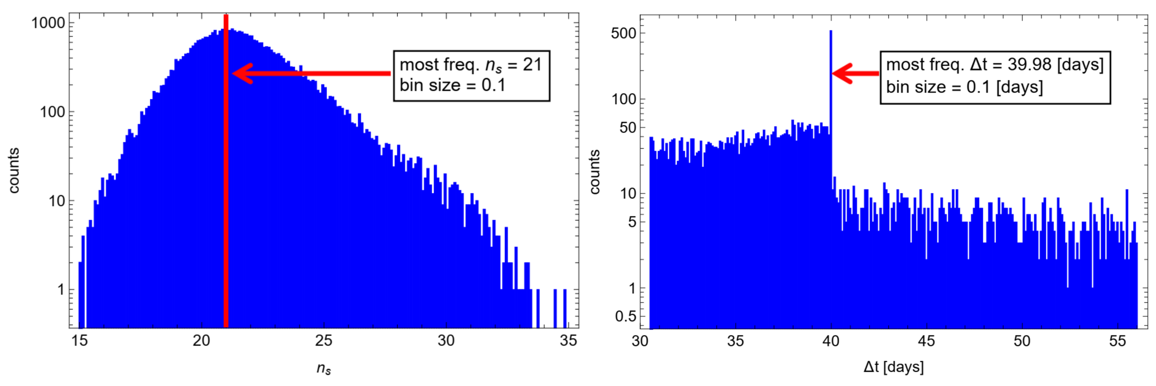
Monte Carlo test

- Randomly generate many sample maps with background and signal events (distributed around an assumed source position and within assumed flares duration ΔT_{flare})



- **For each map we obtain:**
  - estimator n_s of number of signal events N_s
  - estimator ΔT of the flares duration ΔT_{flare}
  - value of the test statistic $\widetilde{TS}(M_{opt})$

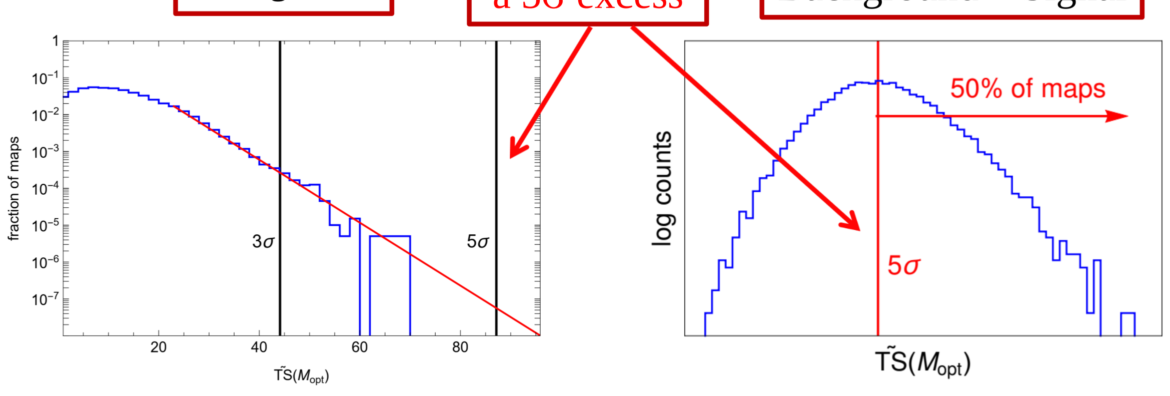
- 3 flares, flares duration 20, 10 and 10 days, number of signal events N_s = 20



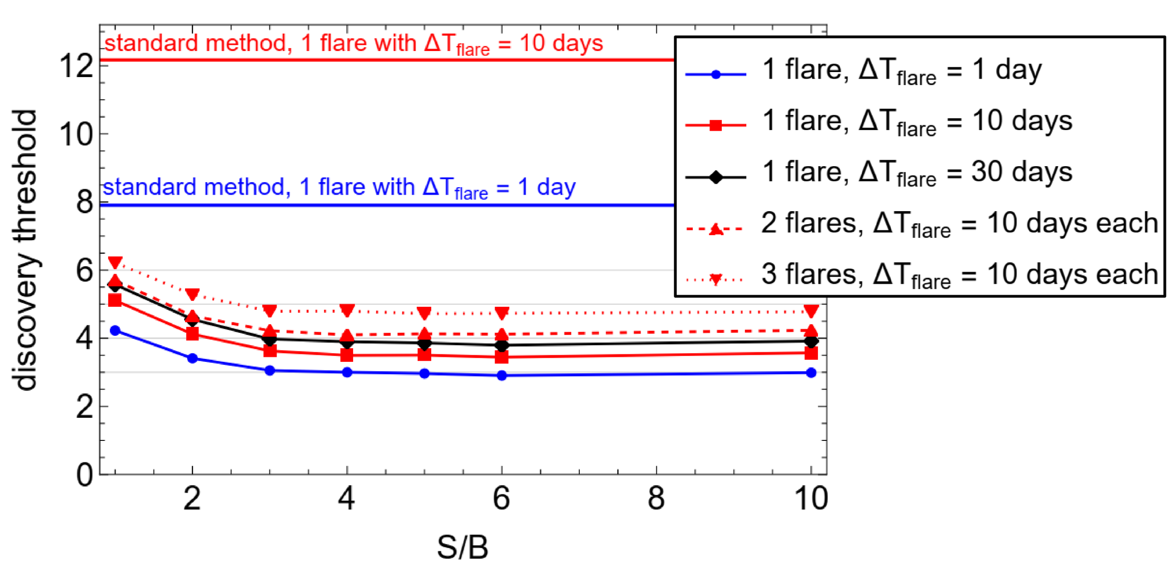
- Number of signal events N_s = 20 and total flares duration ΔT_{flare} = 40 days are recovered.

Discovery potential of the stacking space-time clustering method with S_b photon tag

- Discovery potential tells us how many signal events are needed to claim discovery of a cluster of events in data. It can be used to compare different methods.
- **Definition:** The discovery threshold is the number of signal events required to achieve a p-value less than 2.87 x 10^{-7} (one-sided 5σ) in 50% of the maps (i.e. the number of signal events for which median of test statistic $\widetilde{TS}(M_{opt})$ distribution is at 5σ threshold).



- Discovery thresholds vs threshold for signal-like events S/B



- Only a few events are needed to detect flares at a higher S/B threshold for signal-like events.

Summary

- Advantages of the stacking method:
  - it is faster than the standard method
  - more sensitive to weak flares of any shapes
  - able to detect multiple flares.
- The stacking method is able to recover the number of signal events and flares duration with small uncertainty.
- **The stacking method with S_b photon tag requires only 3 to 5 events to discover photon flares!**

- J. Braun et al. Astropart. Phys. 29 (2008) 299
- D. Góra et al. Astropart. Phys. 35 (2011) 201
- Universe 8 (2022) 579, astro-ph.HE/2210.12959
- G. Ros et al.. Astropart. Phys. 35 (2011) 140

his work was partly funded by National Science Centre, Grants No. 2020/39/B/ST9/01398 and 2022/45/B/ST9/02163, and by Ministry of Science and Higher Education, Grant No. 2022/WK/12